

OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/836,561

DATE: 06/13/2001

TIME: 16:42:37

Input Set : N:\Cr3\RULE60\09836561.txt

Output Set: N:\CRF3\06132001\I836561.raw

ENTERED

SEQUENCE LISTING

```

C--> 5 (1) GENERAL INFORMATION:
      7 (i) APPLICANT: Bandman, Olga
      8 Corley, Neil C.
      9 Guegler, Karl J.
C--> 11 (ii) TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
      13 (iii) NUMBER OF SEQUENCES: 6
      15 (iv) CORRESPONDENCE ADDRESS:
      16 (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
      17 (B) STREET: 3174 Porter Drive
      18 (C) CITY: Palo Alto
      19 (D) STATE: CA
      20 (E) COUNTRY: USA
      21 (F) ZIP: 94304
      23 (v) COMPUTER READABLE FORM:
      24 (A) MEDIUM TYPE: Diskette
      25 (B) COMPUTER: IBM Compatible
      26 (C) OPERATING SYSTEM: DOS
      27 (D) SOFTWARE: FastSEQ for Windows Version 2.0
      29 (vi) CURRENT APPLICATION DATA:
C--> 30 (A) APPLICATION NUMBER: US/09/836,561
C--> 31 (B) FILING DATE: 16-Apr-2001
      32 (C) CLASSIFICATION:
      34 (vii) PRIOR APPLICATION DATA:
      35 (A) APPLICATION NUMBER: 09/212,168
      36 (B) FILING DATE:
      38 (viii) ATTORNEY/AGENT INFORMATION:
      39 (A) NAME: Billings, Lucy J.
      40 (B) REGISTRATION NUMBER: 36,749
      41 (C) REFERENCE/DOCKET NUMBER: PF-0333 US
      43 (ix) TELECOMMUNICATION INFORMATION:
      44 (A) TELEPHONE: 415-855-0555
      45 (B) TELEFAX: 415-845-4166
      46 (C) TELEX:
      49 (2) INFORMATION FOR SEQ ID NO: 1:
      51 (i) SEQUENCE CHARACTERISTICS:
      52 (A) LENGTH: 448 amino acids
      53 (B) TYPE: amino acid
      54 (C) STRANDEDNESS: single
      55 (D) TOPOLOGY: linear
      57 (vii) IMMEDIATE SOURCE:
      58 (A) LIBRARY: CORNNOT01
      59 (B) CLONE: 45517
      61 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
      63 Met Pro Gly Ile Lys Arg Ile Leu Thr Val Thr Ile Leu Ala Leu Cys
      64 1 5 10 15
      65 Leu Pro Ser Pro Gly Asn Ala Gln Ala Gln Cys Thr Asn Gly Phe Asp

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/836,561

DATE: 06/13/2001

TIME: 16:42:37

Input Set : N:\Cr3\RULE60\09836561.txt

Output Set: N:\CRF3\06132001\I836561.raw

```

66          20          25          30
67 Leu Asp Arg Gln Ser Gly Gln Cys Leu Asp Ile Asp Glu Cys Arg Thr
68          35          40          45
69 Ile Pro Glu Ala Cys Arg Gly Asp Met Met Cys Val Asn Gln Asn Gly
70          50          55          60
71 Gly Tyr Leu Cys Ile Pro Arg Thr Asn Pro Val Tyr Arg Gly Pro Tyr
72          65          70          75          80
73 Ser Asn Pro Tyr Ser Thr Pro Tyr Ser Gly Pro Tyr Pro Ala Ala Ala
74          85          90          95
75 Pro Pro Leu Ser Ala Pro Asn Tyr Pro Thr Ile Ser Arg Pro Leu Ile
76          100         105         110
77 Cys Arg Phe Gly Tyr Gln Met Asp Glu Ser Asn Gln Cys Val Asp Val
78          115         120         125
79 Asp Glu Cys Ala Thr Asp Ser His Gln Cys Asn Pro Thr Gln Ile Cys
80          130         135         140
81 Ile Asn Thr Glu Gly Gly Tyr Thr Cys Ser Cys Thr Asp Gly Tyr Trp
82          145         150         155         160
83 Leu Leu Glu Gly Gln Cys Leu Asp Ile Asp Glu Cys Arg Tyr Gly Tyr
84          165         170         175
85 Cys Gln Gln Leu Cys Ala Asn Val Pro Gly Ser Tyr Ser Cys Thr Cys
86          180         185         190
87 Asn Pro Gly Phe Thr Leu Asn Glu Asp Gly Arg Ser Cys Gln Asp Val
88          195         200         205
89 Asn Glu Cys Ala Thr Glu Asn Pro Cys Val Gln Thr Cys Val Asn Thr
90          210         215         220
91 Tyr Gly Ser Phe Ile Cys Arg Cys Asp Pro Gly Tyr Glu Leu Glu Glu
92          225         230         235         240
93 Asp Gly Val His Cys Ser Asp Met Asp Glu Cys Ser Phe Ser Glu Phe
94          245         250         255
95 Leu Cys Gln His Glu Cys Val Asn Gln Pro Gly Thr Tyr Phe Cys Ser
96          260         265         270
97 Cys Pro Pro Gly Tyr Ile Leu Leu Asp Asp Asn Arg Ser Cys Gln Asp
98          275         280         285
99 Ile Asn Glu Cys Glu His Arg Asn His Thr Cys Asn Leu Gln Gln Thr
100         290         295         300
101 Cys Tyr Asn Leu Gln Gly Gly Phe Lys Cys Ile Asp Pro Ile Arg Cys
102          305         310         315         320
103 Glu Glu Pro Tyr Leu Arg Ile Ser Asp Asn Arg Cys Met Cys Pro Ala
104          325         330         335
105 Glu Asn Pro Gly Cys Arg Asp Gln Pro Phe Thr Ile Leu Tyr Arg Asp
106          340         345         350
107 Met Asp Val Val Ser Gly Arg Ser Val Pro Ala Asp Ile Phe Gln Met
108          355         360         365
109 Gln Ala Thr Thr Arg Tyr Pro Gly Ala Tyr Tyr Ile Phe Gln Ile Lys
110          370         375         380
111 Ser Gly Asn Glu Gly Arg Glu Phe Tyr Met Arg Gln Thr Gly Pro Ile
112          385         390         395         400
113 Ser Ala Thr Leu Val Met Thr Arg Pro Ile Lys Gly Pro Arg Glu Ile
114          405         410         415

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/836,561

DATE: 06/13/2001

TIME: 16:42:37

Input Set : N:\Cr3\RULE60\09836561.txt

Output Set: N:\CRF3\06132001\I836561.raw

115 Gln Leu Asp Leu Glu Met Ile Thr Val Asn Thr Val Ile Asn Phe Arg
 116 420 425 430
 117 Gly Ser Ser Val Ile Arg Leu Arg Ile Tyr Val Ser Gln Tyr Pro Phe
 118 435 440 445

120 (2) INFORMATION FOR SEQ ID NO: 2:

122 (i) SEQUENCE CHARACTERISTICS:

123 (A) LENGTH: 2550 base pairs

124 (B) TYPE: nucleic acid

125 (C) STRANDEDNESS: single

126 (D) TOPOLOGY: linear

128 (vii) IMMEDIATE SOURCE:

129 (A) LIBRARY: CORNNOT01

130 (B) CLONE: 45517

132 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

134	CCAAGATTGT	TGTGAGGAGT	CTAGCCAGTT	GGTGAGCGCT	GTAATCTGAA	CCAGCTGTGT	60
135	CCAGACTGAG	GCCCCATTTG	CATTATTTAA	CATACTTAGA	AAATGAAGTG	TTCATTTTAA	120
136	ACATTCCCTC	TCCAATTGGT	TTAATGCTGA	ATTACTGAAG	AGGGCTAAGC	AAAACCAGGT	180
137	GCTTGCGCTG	AGGGCTCTGC	AGTGGCTGGG	AGGACCCCGG	CGCTCTCCCC	GTGTCTCTCT	240
138	CACGACTCGC	TCCGCCCTC	TGGAATAAAA	CACCCGCGAG	CCCCGAGGGC	CCAGAGGAGG	300
139	CCGACGTGCC	CGAGCTCCTC	CGGGGGTCCC	GCCCGCGAGC	TTTCTTCTCG	CCTTCGCATC	360
140	TCCTCCTCGC	GCGTCTTGGA	CATGCCAGGA	ATAAAAAGGA	TACTCACTGT	TACCATTCTG	420
141	GCTCTCTGTC	TTCCAAGCCC	TGGGAATGCA	CAGGCACAGT	GCACGAATGG	CTTTGACCTG	480
142	GATCGCCAGT	CAGGACAGTG	TTTAGATATT	GATGAATGCC	GAACCATCCC	CGAGGCCTGC	540
143	CGAGGAGACA	TGATGTGTGT	TAACCAAAAT	GGCGGGTATT	TATGCATTCC	CCGGACAAAC	600
144	CCTGTGTATC	GAGGGCCCTA	CTCGAACCCC	TACTCGACCC	CCTACTCAGG	TCCGTACCCA	660
145	GCAGCTGCCC	CACCACTCTC	AGCTCCAAAC	TATCCCACGA	TCTCCAGGCC	TCTTATATGC	720
146	CGCTTTGGAT	ACCAGATGGA	TGAAAGCAAC	CAATGTGTGG	ATGTGGACGA	GTGTGCAACA	780
147	GATTCCCACC	AGTGCAACCC	CACCCAGATC	TGCATCAATA	CTGAAGGCGG	GTACACCTGC	840
148	TCCTGCACCG	ACGGATATTG	GCTTCTGGAA	GGCCAGTGCT	TAGACATTGA	TGAATGTCGC	900
149	TATGGTACT	GCCAGCAGCT	CTGTGCGAAT	GTTCTTGAT	CCTATTCTTG	TACATGCAAC	960
150	CCTGGTTTAA	CCCTCAATGA	GGATGGAAGG	TCTTGCCAAG	ATGTGAACGA	GTGTGCCACC	1020
151	GAGAACCCCT	GCGTGCAAAC	CTGCGTCAAC	ACCTACGGCT	CTTTCATCTG	CCGCTGTGAC	1080
152	CCAGGATATG	AACTTGAGGA	AGATGGCGTT	CATTGCAGTG	ATATGGACGA	GTGCAGCTTC	1140
153	TCTGAGTTCC	TCTGCCAACA	TGAGTGTGTG	AACCAGCCCG	GCACATACTT	CTGCTCCTGC	1200
154	CCTCCAGGCT	ACATCCTGCT	GGATGACAAC	CGAAGCTGCC	AAGACATCAA	CGAATGTGAG	1260
155	CACAGGAACC	ACACGTGCAA	CCTGCAGCAG	ACGTGCTACA	ATTTACAAGG	GGGCTTCAAA	1320
156	TGCATCGACC	CCATCCGCTG	TGAGGAGCCT	TATCTGAGGA	TCAGTGATAA	CCGCTGTATG	1380
157	TGTCTTGCTG	AGAACCCTGG	CTGCAGAGAC	CAGCCCTTTA	CCATCTTGTA	CCGGGACATG	1440
158	GACGTGGTGT	CAGGACGCTC	CGTTCCCGCT	GACATCTTCC	AAATGCAAGC	CACGACCCGC	1500
159	TACCCTGGGG	CCTATTACAT	TTTCCAGATC	AAATCTGGGA	ATGAGGGCAG	AGAAATTTAC	1560
160	ATGCGGCAAA	CGGGCCCAT	CAGTGCCACC	CTGGTGATGA	CACGCCCCAT	CAAAGGGCCC	1620
161	CGGGAAATCC	AGCTGGACTT	GGAAATGATC	ACTGTCAACA	CTGTCATCAA	CTTCAGAGGC	1680
162	AGCTCCGTGA	TCCGACTGCG	GATATATGTG	TCGCAGTACC	CATTCTGAGC	CTCGGGCTGG	1740
163	AGCCTCCGAC	GCTGCCTCTC	ATTGGCACCA	AGGGACAGGA	GAAGAGAGGA	AATAACAGAG	1800
164	AGAATGAGAG	CGACACAGAC	GTTAGGCATT	TCCTGCTGAA	CGTTTCCCCG	AAGAGTCAGC	1860
165	CCCGACTTCC	TGACTCTCAC	CTGTACTATT	GCAGACCTGT	CACCCCTGCAG	GACTTGCCAC	1920
166	CCCCAGTTCC	TATGACACAG	TTATCAAAAA	GTATTATCAT	TGCTCCCCTG	ATAGAAGATT	1980
167	GTTGGTGAAT	TTTCAAGGCC	TTCAGTTTAT	TTCCACTATT	TTCAAAGAAA	ATAGATTAGG	2040
168	TTTGCGGGGG	TCTGAGTCTA	TGTTCAAAGA	CTGTGAACAG	CTTGCTGTCA	CTTCTTCACC	2100

RAW SEQUENCE LISTING

DATE: 06/13/2001

PATENT APPLICATION: US/09/836,561

TIME: 16:42:37

Input Set : N:\Cr3\RULE60\09836561.txt

Output Set: N:\CRF3\06132001\I836561.raw

```

169 TCTTCCACTC CTTCTCTCAC TGTGTTACTG CTTTGCAAAG ACCCGGGAGC TGGCGGGGAA 2160
170 CCCTGGGAGT AGCTAGTTTG CTTTTTGCGT ACACAGAGAA GGCTATGTAA ACAAACCACA 2220
171 GCAGGATCGA AGGGTTTTTA GAGAATGTGT TTCAAACCA TGCCTGGTAT TTTCAACCAT 2280
172 AAAAGAAGTT TCAGTTGTCC TTAAATTTGT ATAACGGTTT AATTCTGTCT TGTTCATTTT 2340
173 GAGTATTTTT AAAAAATATG TCGTAGAATT CCTTCGAAAG GCCTTCAGAC ACATGCTATG 2400
174 TTCTGTCTTC CCAAACCCAG TCTCCTCTCC ATTTTAGCCC AGTGTCTTCT TTGAGGACCC 2460
175 CTTAATCTTG CTTTCTTTAG AATTTTACC CAATTGGATT GGAATGCAGA GGTCTCCAAA 2520
176 CTGATTAAAT ATTTGAAGAG AAAAAAAAAA 2550

```

178 (2) INFORMATION FOR SEQ ID NO: 3:

180 (i) SEQUENCE CHARACTERISTICS:

181 (A) LENGTH: 540 amino acids

182 (B) TYPE: amino acid

183 (C) STRANDEDNESS: single

184 (D) TOPOLOGY: linear

186 (vii) IMMEDIATE SOURCE:

187 (A) LIBRARY: BRAITUT13

188 (B) CLONE: 1621777

190 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```

192 Met Gly Thr Thr Ala Arg Ala Ala Leu Val Leu Thr Tyr Leu Ala Val
193 1 5 10 15
194 Ala Ser Ala Ala Ser Glu Gly Gly Phe Thr Ala Thr Gly Gln Arg Gln
195 20 25 30
196 Leu Arg Pro Glu His Phe Gln Glu Val Gly Tyr Ala Ala Pro Pro Ser
197 35 40 45
198 Pro Pro Leu Ser Arg Ser Leu Pro Met Asp His Pro Asp Ser Ser Gln
199 50 55 60
200 His Gly Pro Pro Phe Glu Gly Gln Ser Gln Val Gln Pro Pro Pro Ser
201 65 70 75 80
202 Gln Glu Ala Thr Pro Leu Gln Gln Glu Lys Leu Leu Pro Ala Gln Leu
203 85 90 95
204 Pro Ala Glu Lys Glu Val Gly Pro Pro Leu Pro Gln Glu Ala Val Pro
205 100 105 110
206 Leu Gln Lys Glu Leu Pro Ser Leu Gln His Pro Asn Glu Gln Lys Glu
207 115 120 125
208 Gly Met Pro Ala Pro Phe Gly Asp Gln Ser His Pro Glu Pro Glu Ser
209 130 135 140
210 Trp Asn Ala Ala Gln His Cys Gln Gln Asp Arg Ser Gln Gly Gly Trp
211 145 150 155 160
212 Gly His Arg Leu Asp Gly Phe Pro Pro Gly Arg Pro Ser Pro Asp Asn
213 165 170 175
214 Leu Asn Gln Ile Cys Leu Pro Asn Arg Gln His Val Val Tyr Gly Pro
215 180 185 190
216 Trp Asn Leu Pro Gln Ser Ser Tyr Ser His Leu Thr Arg Gln Gly Glu
217 195 200 205
218 Thr Leu Asn Phe Leu Glu Ile Gly Tyr Ser Arg Cys Cys His Cys Arg
219 210 215 220
220 Ser His Thr Asn Arg Leu Glu Cys Ala Lys Leu Val Trp Glu Glu Ala
221 225 230 235 240
222 Met Ser Arg Phe Cys Glu Ala Glu Phe Ser Val Lys Thr Arg Pro His

```

RAW SEQUENCE LISTING

DATE: 06/13/2001

PATENT APPLICATION: US/09/836,561

TIME: 16:42:37

Input Set : N:\Cr3\RULE60\09836561.txt

Output Set: N:\CRF3\06132001\I836561.raw

```

223          245          250          255
224 Trp Cys Cys Thr Arg Gln Gly Glu Ala Arg Phe Ser Cys Phe Gln Glu
225          260          265          270
226 Glu Ala Pro Gln Pro His Tyr Gln Leu Arg Ala Cys Pro Ser His Gln
227          275          280          285
228 Pro Asp Ile Ser Ser Gly Leu Glu Leu Pro Phe Pro Pro Gly Val Pro
229          290          295          300
230 Thr Leu Asp Asn Ile Lys Asn Ile Cys His Leu Arg Arg Phe Arg Ser
231          305          310          315          320
232 Val Pro Arg Asn Leu Pro Ala Thr Asp Pro Leu Gln Arg Glu Leu Leu
233          325          330          335
234 Ala Leu Ile Gln Leu Glu Arg Glu Phe Gln Arg Cys Cys Arg Gln Gly
235          340          345          350
236 Asn Asn His Thr Cys Thr Trp Lys Ala Trp Glu Asp Thr Leu Asp Lys
237          355          360          365
238 Tyr Cys Asp Arg Glu Tyr Ala Val Lys Thr His His His Leu Cys Cys
239          370          375          380
240 Arg His Pro Pro Ser Pro Thr Arg Asp Glu Cys Phe Ala Arg Arg Ala
241          385          390          395          400
242 Pro Tyr Pro Asn Tyr Asp Arg Asp Ile Leu Thr Ile Asp Ile Gly Arg
243          405          410          415
244 Val Thr Pro Asn Leu Met Gly His Leu Cys Gly Asn Gln Arg Val Leu
245          420          425          430
246 Thr Lys His Lys His Ile Pro Gly Leu Ile His Asn Met Thr Ala Arg
247          435          440          445
248 Cys Cys Asp Leu Pro Phe Pro Glu Gln Ala Cys Cys Ala Glu Glu Glu
249          450          455          460
250 Lys Leu Thr Phe Ile Asn Asp Leu Cys Gly Pro Arg Arg Asn Ile Trp
251          465          470          475          480
252 Arg Asp Pro Ala Leu Cys Cys Tyr Leu Ser Pro Gly Asp Glu Gln Val
253          485          490          495
254 Asn Cys Phe Asn Ile Asn Tyr Leu Arg Asn Val Ala Leu Val Ser Gly
255          500          505          510
256 Asp Thr Glu Asn Ala Lys Gly Gln Gly Glu Gln Gly Ser Thr Gly Gly
257          515          520          525
258 Thr Asn Ile Ser Ser Thr Ser Glu Pro Lys Glu Glu
259          530          535          540
261 (2) INFORMATION FOR SEQ ID NO: 4:
262     (i) SEQUENCE CHARACTERISTICS:
263         (A) LENGTH: 1899 base pairs
264         (B) TYPE: nucleic acid
265         (C) STRANDEDNESS: single
266         (D) TOPOLOGY: linear
267     (vii) IMMEDIATE SOURCE:
268         (A) LIBRARY: BRAITUT13
269         (B) CLONE: 162177
270     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
271 TGGGTGCAAG CTCACAACCG TAACAGCCAC CAGACAAGCT TCAGTGGCCG GCCCTTCACA      60
272 TCCAGACTTG CCTGAGAGGA CCCACCTCTG AGTGTCCAGT GGTGAGTTGC CCCAGGATGG      120

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/836,561

DATE: 06/13/2001

TIME: 16:42:38

Input Set : N:\Crf3\RULE60\09836561.txt

Output Set: N:\CRF3\06132001\I836561.raw

L:5 M:220 C: Keyword misspelled or invalid format, [(1) GENERAL INFORMATION:]
L:11 M:220 C: Keyword misspelled or invalid format, [(ii) TITLE OF INVENTION:]
L:30 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:31 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]